Journal of Risk and Uncertainty, 8, 21–32 (1999)


# Calibrating Hypothetical Willingness to Pay Responses

MAGNUS JOHANNESSON
*Department of Economics, Stockholm School of Economics, Stockholm, Sweden*

GLENN C. BLOMQUIST
*Department of Economics and Martin School of Public Policy and Administration, University of Kentucky, Lexington, USA*

KAREN BLUMENSCHEIN
*College of Pharmacy, University of Kentucky, Lexington, USA*

PER-OLOV JOHANSSON
*Department of Economics, Stockholm School of Economics, Stockholm, Sweden*

BENGT LILJAS
*Department of Economics, Lund University, Lund, Sweden*

RICHARD M. O'CONOR
*Consultant, Park City, Utah, USA*

## *Abstract*

Experimental data comparing hypothetical and real dichotomous choice responses for two different goods were used to estimate a statistical bias function to calibrate the hypothetical yes responses. The probability that a hypothetical yes response would be a real yes response was estimated as a function of the individual's self-assessed certainty of the hypothetical yes response (assessed on a 0–10 scale) and a variable representing the price level. Without calibration the hypothetical yes responses significantly exceeded the proportion of real yes responses, but after calibration the null hypothesis of no difference between hypothetical and real responses could not be rejected in any of the experiments.

**Key words:** contingent valuation, hypothetical bias, willingness to pay, experiments, calibration

**JEL Classification:** D61, D89, C91

The contingent valuation method has been developed to measure the willingness to pay for environmental changes and other non-market goods such as health and safety changes. Individuals are asked about their hypothetical willingness to pay for a defined good (Mitchell and Carson, 1989). In the most commonly used elicitation approach, the dichotomous choice approach, individuals accept or reject only one price and opportunity to purchase the defined good. By varying the price in different subsamples it is possible to derive the demand curve and estimate the mean willingness to pay (Hanemann, 1984). The hypothetical dichotomous choice question often is framed in terms of a vote on a referendum, to increase realism (Mitchell and Carson, 1989).

The use of the contingent valuation method is controversial among economists (Hanemann, 1994; Diamond and Hausman, 1994). The nucleus of the controversy is the extent to which hypothetical choices in the contingent valuation method correspond to real economic choices. The extent to which hypothetical choices mimic real choices is an old controversy in economics. Already in 1942 Wallis and Friedman (1942) criticised the use of hypothetical choices in experiments, and it has been debated ever since (Kagel and Roth, 1995; Thaler, 1987).

Until recently, however, relatively little work has been carried out directly comparing hypothetical and real willingness to pay. Cummings, Harrison, and Rutström (1995) recently presented data from an experiment comparing dichotomous choice contingent valuation choices with real choices. Three experiments were carried out on three different consumer goods (an electric juice-maker, a calculator, and a box of chocolates). For each good the proportion of hypothetical yes responses significantly exceeded the proportion of real yes responses. The overestimation of the dichotomous choice approach noted by Cummings, Harrison, and Rutström (1995) was also confirmed in two other experiments on private goods by Johannesson, Liljas, and Johansson (1998) and Blumenschein et al. (1998). In another experiment Cummings et al. (1997) compared a hypothetical referenda with a real referenda, and found that the proportion of hypothetical yes responses significantly exceeded the proportion of real yes responses. Thus, the framing of the hypothetical question in terms of a referendum did not remove the overestimation problem. In early work by Bishop and Heberlein (1979) and in more recent work by Nape et al. (1995), dichotomous choice contingent valuation questions to measure willingness to accept have also been shown to lead to similar problems of overestimation.

The case can be made then that the dichotomous choice contingent valuation approach is associated with a general overestimation problem. The question arisis as to whether it is possible to calibrate hypothetical dichotomous choice responses to better correspond to real decisions. In this paper we present a method for identifying the subset for which hypothetical yes responses represent real yes responses and for calibrating the full set of hypothetical responses. This paper builds on recent related work. Blackburn, Harrison, and Rutström (1994) tried to predict the true yes responses from the Cummings, Harrison and Rutström (1995) data from 1992, based on the socio-economic characteristics of the respondents. Although the socio-economic characteristics had limited explanatory power, the estimated statistical bias functions had some ability to correct for the overestimation in the hypothetical responses. In two experiments Johannesson, Liljas and Johansson (1998) and Blumenschein et al. (1998) tested the hypothesis that only definitely sure yes responses, identified in a follow-up question, correspond to real yes responses. In the Johannesson, Liljas, and Johansson (1998) experiment the definitely sure yes responses significantly underestimated the real yes responses whereas in the Blumenshein et al. (1998) experiment the null hypothesis of no difference between definitely sure yes responses and real yes responses could not be rejected. A related approach was also used in a recent study by Champ et al. (1997), which compared hypothetical dichotomous choice questions about donating a specified amount to a public good with actual donations to the public good. They assessed the certainty of the hypothetical donation responses on a 1-10 scale from very uncertain to very certain. They found that hypothetical donations significantly exceeded real donations, but that there was no significant difference if only subjects that were very certain of their yes responses were counted as real yes responses.

A similar certainty scale as used by Champ et al. (1997) was also included in the experiments by Johannesson and Liljas, (1998) and Blumenschein et al (1998). Data was collected about the degree of certainty of the hypothetical yes responses on a scale between 0 (very unsure) and 10 (very sure). The specific purpose of this paper is to estimate a statistical bias function with the certainty scale as an explanatory variable, and to test if this bias function can correct for the dichotomous choice overestimation observed in the Johannesson and Liljas (1998) and Blumenschein et al. (1998) experiments. Below we first outline the methods used to estimate and test the statistical bias function. The results are then presented and the paper ends with some concluding remarks.

## 1. Methods

We used data from two experiments that compared hypothetical and real yes responses for private goods (Johannesson and Liljas, 1998; Blumenschein et al., 1998).[1] The good used in the experiment reported by Johannesson, Liljas, and Johansson (1998) was a box of Belgian chocolates (500 g), and the experiment was carried out on a group of 242 undergraduate students at Lund University in Sweden. Three different prices were used (SEK 20, SEK 30 and SEK 50; $1 = SEK 8.0) and the proportion of hypothetical and real yes responses was compared both between and within samples. The good used in the experiment reported by Blumenschein et al. (1998) was a pair of sunglasses and 133 undergraduate students at the University of Kentucky College of Business and Economics were included in the experiment. Two different prices were used ($1 and $5) and the proportion of hypothetical and real yes responses were compared both between and within samples.[2] In both experiments the proportion of hypothetical yes responses was significantly higher than the proportion of real yes responses.

To estimate a statistical bias function, information is needed about both hypothetical and real responses from the same individuals. Therefore, we used only the data from the within samples comparisons in the two experiments, i.e. the respondents that received a hypothetical dichotomous choice question followed by a real dichotomous choice question. The proportion of real yes responses did not differ significantly between the sample that received the real dichotomous choice question after the hypothetical dichotomous choice question and the sample that received only the real dichotomous choice question in any of the experiments. There was thus no evidence of an ordering effect in the experiments, which is also in line with the Cummings, Harrison, and Rutström (1995) results.

To increase the number of observations on the sunglasses good we also carried out an additional within sample experiment. Two different prices ($1 and $3) were used and 84 students at the University of Kentucky College of Pharmacy were included in the experiment.[3] These responses were pooled with the within sample data from the Blumenschein et al. (1998) study using the same good.[4]

Following Blackburn, Harrison, and Rutström (1994) the within samples responses in dichotomous choice experiments can be divided into three categories. The first category is a hypothetical yes response followed by a real yes response (yes-yes), the second category is a hypothetical yes response followed by a real no response (yes-no), and the third category is a hypothetical no response followed by a real no response (no-no).[5] The number and proportion of individuals in each response category in the experiments are shown in Table 1. Since a hypothetical no response was always followed by a real no response in the experiments no calibration is needed for hypothetical no responses. The interesting issue from a calibration viewpoint is to separate the yes-yes responses from the yes-no responses. To try and do this we estimated a statistical bias function for the hypothetical yes responses. In the chocolate experiment 64 individuals answered yes to the hypothetical dichotomous choice question, and of these individuals 48 (75%) responded yes to the real dichotomous choice question and made the purchase. In the sunglasses experiments 35 individuals answered yes to the hypothetical dichotomous choice question, and of these individuals 13 (37%) responded yes to the real dichotomous choice question and made the purchase.

A probit regression analysis was used to estimate the probability of a yes-yes response among the 99 individuals that answered yes to the hypothetical dichotomous choice question (Greene, 1993). Our primary hypothesis was to test if the certainty of the yes response could be used to calibrate the dichotomous choice yes responses. The certainty of the hypothetical yes responses was assessed on a visual

*Table 1.* Within sample response patterns in the experiments.

| | Good | | | |
|---|---|---|---|---|
| | Chocolates | | Sunglasses | |
| Response category | n | % | n | % |
| Hypothetical yes and real yes (yes-yes) | 48 | 39.0 | 13 | 8.7 |
| Hypothetical yes and real no (yes-no) | 16 | 13.0 | 22 | 14.8 |
| Hypothetical no and real no (no-no) | 59 | 48.0 | 114 | 76.5 |
| Total | 123 | 100.00 | 149 | 100.0 |

analogue scale between 0 (very unsure) and 10 (very sure) in the experiments.[6,7] The value on the certainty scale was used as an explanatory variable in the regression analysis. Since it is possible that the bias differs at different price levels, we also included the proportion of hypothetical yes responses at the price faced by the respondent as an explanatory variable. We used the percentage of yes responses rather than the bid level as such, since we only wanted to use variables that can be used to predict the behaviour across different goods (and what is a high price for one good may be a low price for another good). In addition we also included the socio-economic variables age and sex as explanatory variables. The mean and standard deviation of the explanatory variables used in the regression analysis are shown in Table 2

We estimated the probit regression equation for the pooled sample of 99 respondents. For a statistical bias function to be useful it has to be stable across different goods. This can be tested by testing the assumption of pooling the samples for the two goods. We tested the pooling assumption by testing if a dummy variable for the sample was statistically significant and by testing for structural differences between the regression equations for the two samples. The estimated statistical bias function was used to calibrate the hypothetical yes responses in the two experiments. If the predicted probability of a yes-yes response was above 0.5 it was counted as a calibrated yes response and otherwise it was counted as a calibrated no response. A nonparametric sign test was used to test if the proportion of calibrated hypothetical yes responses differed from the real proportion of yes responses in the experiments (Newbold, 1991).

## 2. Results

The probit regression equations of the yes-yes and yes-no responses of the pooled samples are shown in Table 3.[8] In the first equation with all the explanatory variables the certainty scale variable was highly significant with a positive sign, showing that the probability for a yes-yes answer increased with a higher degree of certainty of the yes answer. Also the variable for the bid level, the proportion of

*Table 2.* Descriptive statistics of the explanatory variables in the regression analysis, n = 99.

| Variable | Mean | STD |
|---|---|---|
| Certainty scale[a] | 7.71 | 2.54 |
| Proportion of hypothetical yes responses | 0.56 | 0.24 |
| Sex[b] | 0.52 | 0.50 |
| Age | 22.73 | 3.00 |

[a]: Measured on a scale between 0 (very unsure) and 10 (very sure).
[b]: 1 = man, 0 = woman.

*Table 3.* Probit regression analysis of the probability of a yes-yes response (Statistical bias functions), standard errors within parentheses.

| Variable | Regression equation 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| Constant | −8.33** | −10.09** | −5.44** | −5.81** |
| | (2.12) | (2.64) | (1.06) | (1.22) |
| Certainty scale | 0.62** | 0.65** | 0.61** | 0.62** |
| | (0.13) | (0.13) | (0.12) | (0.12) |
| Proportion of hypothetical yes responses | 1.68* | 2.51** | 1.92** | 2.28** |
| | (0.68) | (0.94) | (0.66) | (0.88) |
| Sex | 0.50 | 0.64 | | |
| | (0.40) | (0.42) | | |
| Age | 0.12 | 0.15 | | |
| | (0.08) | (0.09) | | |
| Sample dummy | | 0.81 | | 0.32 |
| | | (0.59) | | (0.50) |
| n | 99 | 99 | 99 | 99 |
| Chi-square value | 76.93** | 78.98** | 71.92** | 72.33** |
| Log-likelihood | −27.46 | −26.44 | −29.97 | −29.76 |
| McFadden pseudo-$R^2$ | 0.58 | 0.60 | 0.55 | 0.55 |
| Individual prediction (%) | 86.87 | 87.88 | 84.85 | 84.85 |

**, * = significant at the 1% and 5% level according to a two tailed t-test.

hypothetical yes answers, was significant and had a positive sign. This variable shows that the probability for a yes-yes answer decreased at higher bid levels, and that the bias thus increased at higher bid levels. Neither age or sex were statistically significant. The explanatory power of the equation, measured as the McFadden pseudo-$R^2$, was rather high (0.58) and the equation correctly predicted 87% of the yes-yes and yes-no answers.

To test the assumption of pooling the chocolates and sunglasses samples we added a dummy variable for the sample (equation 2 in Table 3). This dummy variable was, however, not statistically significant at the 10% level ($p = 0.17$). To further test the pooling assumption we tested for structural differences between the chocolates and sunglasses samples. This was done by adding interaction terms between the sample dummy variable and all other variables. A likelihood ratio test (LRT) was then carried out to test if this unrestricted model differed from the restricted model without the sample dummy variable and the interaction term. The LRT statistic for this comparison was 3.02 (5 degrees of freedom), which was not statistically significant at the 10% level. We could thus not reject the null hypothesis of no structural differences between the chocolates and sunglasses samples.[9]

Since the socioeconomic variables (sex and age) were not significant in equation 1, we also estimated the regression equation without age and sex included. In equation 3 in Table 3 this regression equation is shown. The certainty scale variable and the variable for the bid level were significant on the 1% level in this equation. The equation correctly predicted 85% of the yes-yes and yes-no responses and the McFadden pseudo-$R^2$ was 0.55. To test the pooling assumption we again added a sample dummy variable (equation 4 in Table 3), but this variable was not statistically significant ($p = 0.55$). The LRT statistic for structural differences was 1.28 (3 degrees of freedom), which was not statistically significant at the 10% level.

The probit equation with the certainty scale and the variable for the bid level (equation 3 in Table 3) was used to calibrate the proportion of hypothetical yes responses in the chocolates and sunglasses experiments. A non-parametric sign test was used to test if the proportion of calibrated hypothetical yes responses differed from the proportion of real yes responses in the experiments (Newbold, 1991).

In Table 4 we show the number and proportion of hypothetical, calibrated hypothetical and real yes responses in the chocolate experiment at each price level. Before any calibration the proportion of hypothetical yes responses was significantly higher than the real proportion of yes responses at the prices of SEK 20 and SEK 50 and for the overall results (all the prices combined). After calibrating the hypothetical yes answers we could not reject the null hypothesis of no difference between the hypothetical calibrated yes responses and the real yes responses for any of the prices or for the overall proportion of yes responses. The overall proposition of yes responses was exactly the same for the calibrated hypothetical responses as for the real responses (39%).

In Table 5 we show the number and proportion of hypothetical, calibrated hypothetical and real yes responses in the sunglasses experiment for each price level. Before calibration the proportion of hypothetical yes responses was significantly higher than the real proportion of yes responses at the prices of $1 and $3 and for the overall proportion of yes responses. For the calibrated hypothetical yes responses we failed to reject the null hypothesis of no difference between calibrated hypothetical responses and real responses at any of the prices or for the overall results. The total calibrated proportion of yes responses was slightly lower than the real proportion of yes responses (8.1% vs. 8.7%), but the difference was

*Table 4.* Number (%) of hypothetical, calibrated hypothetical and real yes responses in the chocolate experiment.

| Price | Hypothetical | Calibrated hypothetical | Real |
|---|---|---|---|
| SEK 20 | 44/52 (85)* | 39/52 (75) | 37/52 (71) |
| SEK 30 | 12/23 (52) | 8/23 (35) | 10/23 (43) |
| SEK 50 | 8/48 (17)* | 1/48 (2) | 1/48 (2) |
| Total | 64/123 (52)** | 48/123 (39) | 48/123 (39) |

**, * = Significantly different from the real yes responses at the 1% and 5% level according to a sign test.

*Table 5.* Number (%) of hypothetical, calibrated hypothetical and real yes responses in the sunglasses experiment.

| Price | Hypothetical | Calibrated hypothetical | Real |
|---|---|---|---|
| $1 | 25/76 (33)** | 9/76 (12) | 10/76 (13) |
| $3 | 9/41 (22)* | 2/41 (5) | 2/41 (5) |
| $5 | 1/32 (3) | 1/32 (3) | 1/32 (3) |
| Total | 35/149 (23)** | 12/149 (8) | 13/149 (9) |

**, * = Significantly different from the real yes responses at the 1% and 5% level according to a sign test.

not significant. If we combine the proportion of yes answers for both experiments the calibrated proportion of yes responses was nearly identical to the real proportion of yes responses (22.1% vs. 22.4%).

## 3. Concluding Remarks

According to our results the certainty of a hypothetical yes response is a strong predictor of whether or not a hypothetical yes response corresponds to a real yes response. The certainty scale was highly significant in the estimated statistical bias function. This suggests that only individuals who are relatively sure of their hypothetical yes response will buy the good in a real purchasing situation. In addition, the variable for the bid level (the proportion of hypothetical yes responses at the price faced by the individual) was statistically significant, indicating that the bias increases at higher bid levels controlling for the certainty of the yes response. In the estimated statistical bias function, the variable for the bid level affects the certainty scale cut-off value needed for the bias function to predict that a hypothetical yes response is a real yes response.[10]

Without calibration the hypothetical yes responses significantly exceeded the proportion of real yes responses, but after calibration with the estimated bias function the null hypothesis of no difference between hypothetical and real responses could not be rejected at any of the price levels in the two experiments. The overall calibrated proportion of yes responses in the chocolate experiment (39%) was identical to the real proportion of yes responses. In the sunglasses experiment the calibrated proportion of yes responses was slightly lower than the real proportion of yes responses (8.1% vs. 8.7%). Overall in both experiments the proportion of hypothetical yes responses was 36.4%, the proportion of calibrated hypothetical yes responses was 22.1% and the proportion of real yes responses was 22.4%.

Neither of the socio-economic variables (age and sex) were statistically significant in the bias function. The insignificance may in part be due to the homogenous student samples used in the study. Further studies are needed to test if socio-economic variables can further improve the predictive ability of the bias function.

Johannesson, Liljas, and Johansson (1998) and Blumenschein et al. (1998) also tested if the degree of certainty of the hypothetical yes responses could be used to calibrate the responses. They, however, only divided the hypothetical yes responses into two categories of uncertainty ("definitely sure" and "probably sure" yes responses), based on a dichotomous follow-up question. In the Blumenshein et al. (1998) experiment the null hypothesis of no difference between definitely sure yes responses and real yes responses could not be rejected, but in the Johannesson, Liljas, and Johansson (1998) experiment the definitely sure yes responses significantly underestimated the real yes responses. The calibration approach tested in this study thus worked better, since the null hypothesis of no difference between calibrated responses and real responses could not be rejected in any of the experiments.

It is also interesting to compare our results with the recent study by Champ et al. (1997), that assessed the degree of certainty of hypothetical responses on a 1-10 scale from very uncertain to very certain. Champ et al. (1997) found that hypothetical donations significantly exceeded real donations, but that there was no significant difference if only subjects that were very certain of their yes responses (10 on the scale) were counted as real yes responses. We tested the same calibration as used by Champ et al. (1997) on our experimental data, i.e. only hypothetical yes responses with 10 on the certainty scale were counted as real yes responses. That, however, led to an underestimation of the real yes responses in both the chocolate and the sunglasses experiments.[11] In comparing our results to Champ et al. (1997) there are some important differences between the studies that limit the comparability between them. We compared real and hypothetical questions about willingness to pay for a private good. Champ et al. (1997) on the other hand compared real and hypothetical questions about voluntary donations to a public good, where real donations were interpreted as a lower bound on the willingness to pay for the public good. It is possible that the calibration of hypothetical questions differ between voluntary donations and willingness to pay and between private and public goods. Note also, that the issue of calibrating hypothetical willingness to pay responses with the contingent valuation method is not limited to public goods. A growing number of contingent valuation applications are carried out in the health care field on private goods (Johannesson, 1996).

In conclusion, our results suggests that it may be possible to adjust for the overestimation of real willingness to pay with the dichotomous choice contingent valuation approach using a statistical bias function. To apply the estimated bias function in practice to calibrate hypothetical dichotomous choice responses would necessitate the addition of a certainty scale question to dichotomous choice contingent valuation studies. Before the approach is used in practice it is, however, important to further test the stability of the bias function in more experiments. Due to the limited sample size in our estimations (99 individuals) it is important to interpret the results with some caution.

## Acknowledgments

We are grateful to an anonymous referee for helpful comments.

## Notes

1. All data and questionnaires from the experiments are available from the authors.
2. The between samples comparison relates to the comparison of the proportion of yes answers between the group that first received the hypothetical dichotomous choice question and the group that only received the real dichotomous choice question. The within samples comparison relates to the comparison of the proportion of yes answers within the group that received a hypothetical dichotomous choice question followed by a real dichotomous choice question.
3. There were 43 subjects at the $1 price and 41 subjects at the $3 price. The proportion of hypothetical yes responses was 37.2% and the proportion of real yes responses was 18.6% at the $1 price (the difference between real and hypothetical yes responses was significant at the 1% level according to a sign test). At the $3 price, the proportion of hypothetical yes responses was 22% and the proportion of real yes responses was 4.9%. (the difference between real and hypothetical yes responses was significant at the 5% level according to a sign test).
4. We tested the pooling assumption by testing if the proportion of hypothetical or real yes responses differed between the samples. This was done by including a dummy variable for the sample in a probit regression on the probability of a yes answer, controlling for the price level. The sample dummy variable was not significant at the 10% level in the regression equation for hypothetical yes responses or the regression equation for real yes responses. See also note 9.
5. In principle there is also a fourth possible response category: a hypothetical no answer followed by a real yes answer (no-yes). No such answers were observed, however, in the experiments. A hypothetical no answer thus seem to correspond to a real no answer and no calibration is therefore necessary of hypothetical no answers.
6. The phrasing of the certainty scale questions differed slightly between the experiments. In the chocolate experiment the certainty scale question was phrased in the following way (translated from Swedish) "Mark with a cross below, between very unsure (0) and absolutely sure (10), how sure you are that you would buy the box of chocolates here and now at a price of SEK 20." In the sunglasses experiments the certainty scale question was phrased in the following way "Mark with a "x" on the line below, between very unsure (0) and very sure (10) how sure you are that you would buy the sunglasses here and now at a price of $5.00."
7. Champ et al. (1997) included a similar question about the certainty of responses in their comparison of hypothetical and real donation decisions. They measured the degree of certainty on a 1-10 scale from very uncertain to very certain. The contingent valuation study by Li and Mattson (1995) also included a question about the degree of certainty in the dichotomous choice responses (on a scale from 0% confidence to 100% confidence). Li and Mattson (1995), however, used the degree of certainty in an attempt to adjust for random measurement error in dichotomous choice studies rather than to try and adjust for the systematic overestimation of real willingness to pay in dichotomous choice contingent valuation studies.
8. We also tested the probit regression equations in table 3 for heteroskedasticity (Greene, 1993). The LRT test statistic was not significant at the 10% level for any of the regression equations in Table 3.
9. We also tested the assumption of pooling the two sunglasses samples in the estimation of the calibration function. This was done by estimating a calibration function for only the sunglasses good and testing if the calibration function differed between the two sunglasses samples. These tests were carried out in the same way as the tests of pooling the data for the two goods (chocolates and sunglasses). Neither the sample dummy variable or the test for structural differences was significant at the 10% level.

10. The certainty scale cut-off value is the value of the certainty scale where the estimated probit equation predict a 50% probability of a real yes response (a yes-yes answer). At a bid where 90% of the respondents answer yes hypothetically the bias function predicts real yes responses for individuals with a value over 6.09 of the certainty scale. At a bid that 10% of the respondents accept hypothetically a certainty scale value over 8.60 is needed for the bias function to predict a real yes response. If only the certainty scale variable is included as an explanatory variable, the bias function will predict real yes responses for all individuals with a value over 7.26 on the certainty scale independent of the price level. The McFadden pseudo-$R^2$ in a probit equation with only the certainty scale variable included is 0.48 and the individual prediction is 83.84% (the estimated probit regression equation is: $-4.50 + 0.62 *$ certainty scale).
11. Using the calibration used by Champ et al. (1997), the overall proportion of calibrated yes responses was 22.0% and the proportion of real yes responses was 39% in the chocolate experiment ($p < 0.01$). In the sunglasses experiment, the overall proportion of calibrated yes responses was 4.7% and the proportion of real yes responses was 8.7% ($p = 0.07$).

## References


Blackburn, McKinley, Glenn W. Harrison, and Elisabet E. Rutström. (1994). "Statistical Bias Functions and Informative Hypothetical Surveys," *American Journal of Agricultural Economics* 76, 1084–1088.

Blumenschein, Karen, Magnus Johannesson, Glenn C. Blomquist, Bengt Liljas, and Richard M. O'Conor. (1998). "Experimental Results on Expressed Certainty and Hypothetical Bias in Contingent Valuation," *Southern Economic Journal* 65, 169–177.

Bishop, Richard C., and Thomas A. Heberlein. (1979). "Measuring Values of Extramarket Goods: Are Indirect Measures Biased?," *American Journal of Agricultural Economics* 61, 926–930.

Champ, Patricia A., Richard C. Bishop, Thomas C. Brown, and Daniel W. McCollum. (1997). "Using Donation Mechanisms to Value Nonuse Benefits from Public Goods," *Journal of Environmental Economics and Management* 33, 151–162.

Cummings, Ronald G., Steven Elliott, Glenn W. Harrison, and James Murphy. (1997). "Are Hypothetical Referenda Incentive Compatible?" *Journal of Political Economy* 105, 609–621.

Cummings, Ronald G., Glenn W. Harrison, and Elisabet E. Rutström. (1995) "Homegrown Values and Hypothetical Surveys: Is the Dichotomous Choice Approach Incentive-Compatible?," *American Economic Review* 85, 260–266.

Diamond, Peter A., and Jerry A. Hausman. (1994). "Contingent Valuation: Is Some Number Better than No Number?," *Journal of Economic Perspectives* 8, 45–64.

Greene, William H. (1993). *Econometric Analysis*, 2nd Ed. New York: Macmillan.

Hanemann, W. Michael. (1984). "Welfare Evaluations in Contingent Valuation Experiments with Discrete Responses," *American Journal of Agricultural Economics* 66, 332–341.

Hanemann, W. Michael. (1994). "Valuing the Environment Through Contingent Valuation," *Journal of Economic Perspectives* 8, 19–43.

Johannesson, Magnus. (1996). *Theory and Methods of Economic Evaluation in Health Care*. Dordrecht: Kluwer Academic Publishers.

Johannesson, Magnus, Bengt Liljas, and Per-Olov Johansson. (1998). "An Experimental Comparison of Dichotomous Choice Contingent Valuation Questions and Real Purchase Decisions," *Applied Economics* 30, 643–647.

Kagel, John H., and Alvin E. Roth (eds.). (1995). *The Handbook of Experimental Economics*. Princeton, New Jersey: Princeton University Press.

Kriström, Bengt. (1990). "A Non-Parametric Approach to the Estimation of Welfare Measures in Discrete Response Valuation Studies," *Land Economics* 66, 135–139.

Li, Chuan-Zhong, and Leif Mattson. (1995). "Discrete Choice under Preference Uncertainty: An Improved Structural Model for Contingent Valuation," *Journal of Environmental Economics and Management* 28, 256–269.

Mitchell, Robert C., and Richard T. Carson. (1989). *Using Surveys to Value Public Goods*: *The Contingent Valuation Method*. Washington D.C.: Resources for the Future.

Nape, Steven, Peter Frykblom, Glenn W. Harrison, and James C. Lesley. (1995). *Hypothetical Bias and Willingness to Accept*. Economics Working Paper B-95-09, College of Business Administration, University of South Carolina.

Newbold, Paul. (1991). *Statistics for Business and Economics*, Third edition. Englewood Cliffs, New Jersey: Prentice-Hall.

Thaler, Richard. (1987). "The Psychology of Choice and the Assumptions of Economics." in Alvin E. Roth (ed.), *Laboratory Experimentation in Economics*: Six Points of View. Cambridge, Cambridge University Press.

Wallis, W. Allen, and Milton Friedman. (1942). "The Empirical Derivation of Indifference Functions." in O. Lange, F. McIntyre, and T. O. Yntema (eds.), *Studies in Mathematical Economics and Econometrics in Memory of Henry Schultz*. Chicago: University of Chicago Press.